

**Title:** An MSSC-approach to Hungarian classifiers

**Author A's name:**<sup>1</sup> Brigitta R. Schvarcz

**Author A's affiliation:** Bar-Ilan University; Afeka College of Engineering

**Author A's email address:** `bridget.schvarcz@gmail.com`

**Author B's name:** Kata Wohlmuth

**Author B's email address:** `katawohl@gmail.com`

**Abbreviated title:** —

**Keywords:** mass/count distinction, classifiers, numerals, atomicity, lexical semantics, formal semantics

**Abstract:** This paper provides a formal semantic analysis of classifiers in Hungarian. We focus on the puzzle posed by classifier optionality in Hungarian, where most nouns can co-occur with a classifier, but do not have to. Here we show that the presence or absence of classifiers in a numeral expression has semantic consequences. Evidence in support of our analysis comes from nouns that are polysemous and have a physical object and an informational object sense, such as *könyv*, ‘book’, *festmény*, ‘painting’, *magazin*, ‘magazine’. We argue that Hungarian classifiers, such as *darab*, can take count nouns as their complement, and their role is restricting the domain of counting to physically distinct, Maximally Strongly Self-Connected entities (Grimm 2012) in the denotation of the noun they modify.

---

<sup>1</sup>The paper is authored by two first co-authors. The names of the authors should appear in alphabetical order.

# 1 Introduction

Hungarian allows an apparently optional classifier:<sup>2</sup> many nouns can be directly modified by a numeral (1a), but can also be combined with a classifier before being modified by a numeral (1b). In this paper, we refer to the former expressions as numeral expressions (henceforth: Num-expressions), and the latter ones as classifier expressions (henceforth: CL-expressions).<sup>3</sup>

- (1) a. *három könyv*  
three book  
'three books'  
b. *három darab könyv*  
three CL<sub>general</sub> book  
'three books'

This paradigm poses a problem for the traditional categorizations which define two major systems for explaining countability in the nominal domain: (i) a system which uses sortal classifiers, such as in Mandarin Chinese; (ii) a system which uses count nouns that can directly be modified by numerals, such as in English (Greenberg 1974). Given this background, classifier optionality displayed in (1) raises questions about the difference between the Num-expressions and the CL-expressions in Hungarian, as well as about the function of classifiers such as *darab* (CL<sub>general</sub>). The aim of this paper is to give a semantic account of the Hungarian classifier phenomenon, in particular, to analyze the interaction among classifiers, numerals, and nouns that have countable entities in their denotation. We present a novel approach under which the complement of a classifier can be a count noun.

The structure of this paper is as follows: in Section 2, we discuss the mass and count distinction and highlight that Hungarian cannot be easily placed in mass/count typologies, partially because of the classifier optionality exemplified in (1). Section 3 presents novel data that show that the presence or absence of classifiers in cases like (1) has some semantic effect. Then, in Section 4, we give a short overview how our data poses a problem for previous approaches to Hungarian classifier optionality. We set out our analysis of Hungarian classifiers in Section 5, and discuss additional cases our analysis accounts for. Finally, Section 6 concludes the paper.

## 2 The mass/count distinction

### 2.1 Mass/count as a grammatical distinction

Mass and count nouns have distinct grammatical properties that have been observed cross-linguistically. Some of the tests distinguishing mass from count in English include: direct modification by numerals, pluralization, and determiner selection.

First, count nouns directly combine with numerals, resulting in Num-expressions, while mass nouns do not. With numerals, mass nouns require an obligatory mediating element, a classifier or

---

<sup>2</sup>Classifier optionality is a term used to refer to the assumption that in certain constructions the classifier is not necessary, but available, and its presence or absence is not associated with a semantic difference, see e.g. Gil (2013).

<sup>3</sup>The following abbreviations are used in the paper: 1 – first person; ACC – accusative case; CL – classifier; INE – inessive case; PFV – perfective aspect; PL – plural; POSS – possessive; SG – singular; SUBL – sublativ.

a measure phrase, resulting in a CL-expression (2). Second, count nouns can be morphologically marked for plural, while mass nouns cannot (3). Third, determiners are sensitive to the mass/count distinction: some select only count nouns, some select only mass nouns (4).

- (2) a. three books  
b. #three waters  
c. three buckets/litres of water
- (3) a. books  
b. #waters
- (4) a. many/\*much books  
b. \*many/much water

It is widely agreed that the distinction between nouns such as *book* and *water* is a grammatical distinction (Pelletier 1975). Nevertheless, the basis of grammatical countability and how it is encoded in the grammar of particular languages is subject to ongoing debate. While some emphasize that counting relies on the semantic atomic structure of a predicate (Link 1983, Landman 1991, 2003, 2004), others argue that counting is a measure operation under which predicates are evaluated in number of measure units (Krifka 1989, 1995). The former approach treats counting as an operation which applies to sums of atoms and specifies the number of atoms in the denotation of a predicate. The latter approach assumes a covert classifier incorporated into the numeral in order to achieve individuation (Krifka 1995).

Rothstein (2010b, 2017) offers an account that integrates elements of both approaches above: counting is an operation on atoms, and countability is lexically encoded in count nouns. Rothstein assumes the basic denotation of a noun to be mass, and a COUNT operation specifies what counts as a semantic atom in a given context. She further argues that mass/count is a reflection of the distinction between the operations of counting and measuring. According to her approach, counting entities and measuring stuff are two different operations: “counting is putting atomic entities in one-to-one relation with natural numbers [...], measuring is giving value to a quantity on a calibrated dimensional scale” (Rothstein 2010b, p. 5). The main difference between count and mass nouns is in their denotation of sets of entities: count nouns denote sets of entities that can be counted, while mass nouns denote sets of entities that can be measured.

The above approaches are concerned with the ontological basis of the mass/count distinction, which does not play a crucial role in our analysis. We consider the mass/count distinction to be a grammatical phenomenon. In this paper, we only make the ontological commitment that count nouns have atoms in their denotation while mass nouns do not, but we leave aside the question what those atoms exactly are and which atomization operations are involved in accessing them.

There is considerable cross-linguistic variation in how counting and measuring is encoded in the grammar of particular languages. The next section describes the two major types of languages discussed in the mass/count literature.

## 2.2 Mass/count typologies

Greenberg (1974) describes two main types of languages in terms of encoding counting: some languages use a grammatical distinction between count nouns and mass nouns, while others use sortal classifiers. In languages such as English, there are count nouns which can be directly modified by

numerals (5a). In contrast, in languages such as Mandarin Chinese, all nouns are mass and use sortal classifiers upon combining nouns with numerals (5b).

- (5) a. five books  
 b. *wu \*(ben) shu*  
 five itCL<sub>volume</sub> book  
 ‘five books’

Chierchia (1998, 2010) argues that the two systems are in complementary distribution, and that disparate grammatical properties are associated with each system. In typical mass/count languages, direct modification of count nouns by numerals is possible, as shown in (2), and a morphological singular/plural distinction on the noun is marked, see (3); neither is the case in typical classifier languages, see the Mandarin Chinese examples in (6)–(7).

- (6) a. *san \*(ge) pingguo*  
 three CL<sub>general</sub> apple  
 ‘three apples’  
 b. *san \*(ping) shui*  
 three CL<sub>bottle</sub> water  
 ‘three bottles of water’
- (7) a. *yi \*(ben) shu*  
 one CL<sub>volume</sub> book  
 ‘one book’  
 b. *wu \*(ben) shu*  
 five CL<sub>volume</sub> book  
 ‘five books’

In mass/count languages, sortal classifiers are incompatible with count nouns (8), and bare singulars cannot appear in argument position (9); both are possible in classifier languages, see the Mandarin Chinese (7a) and (10).

(8) \*I have one unit/piece of book.

(9) \*I bought book.

(10) *wo mai le shu*  
 I buy PFV book

‘I bought a/the book/books.’

(Schvarcz & Rothstein 2017, p. 184, ex. (2d))

The following table sums up the differences between typical mass/count languages on the one hand, and classifier languages on the other.

	Direct modification by numerals	Plural nominal morphology	Sortal individuating classifiers	Bare singular arguments
Mass/count	✓	✓	✗	✗
Classifier	✗	✗	✓	✓

Table 1: The traits of the two types of languages according to Chierchia (1998, 2010)

### 2.3 Hungarian and the mass/count typologies

Hungarian poses a problem for such a categorization, as it exhibits both genuine mass/count language traits (Schvarcz 2014) and classifier language traits simultaneously. Hungarian allows for the direct modification of nouns by numerals (11) and exhibits a morphological distinction between singular and plural nouns (12) – although nouns are not morphologically marked plural when combined with numerals.

- (11) *három könyv*  
three book  
'three books'
- (12) *könyv/könyv-ek*  
book/book-PL  
'book/books'

Moreover, determiners such as *minden* 'every' are restricted to singular count nouns (13a); while other determiners such as *egy kis* 'some/a bit/a little' are restricted to mass nouns (13b).<sup>4</sup> These traits make Hungarian mass/count language-like.

- (13) a. *minden könyv/#víz*  
every book/water  
'every book/#water'
- b. *egy kis #könyv/víz*  
one little book/water  
'some #book/water'

(Schvarcz 2014, p. 20)

However, Hungarian allows for bare singular arguments (14), and a CL-expression upon combining nouns with numerals, as already shown in (1b) (repeated as (15) below). These traits make the language similar to a classifier language.

- (14) *Könyv-et vettem.*  
book-ACC bought.1SG  
'I bought a book/books.'

---

<sup>4</sup>A reviewer points out that edible nouns like *palacsinta* 'pancake', *körte* 'pear', or *mogyoró* 'peanut' or 'hazelnut' are compatible with *egy kis*, see (ia), indicating mass-like properties, at the same time these nouns can be directly modified by a numeral or the determiner *minden* 'every', as in (ib), suggesting they are count nouns.

- (i) a. *egy kis palacsinta/körte/mogyoró*  
one little pancake/pear/peanut  
'some pancake/pear/peanut'
- b. *három/minden palacsinta/körte/mogyoró*  
three/every pancake/pear/peanut  
'three/every pancake(s)/pear(s)/peanut(s)'

Edible nouns are an area of language where there are consistent mass/count interpretation shifts cross-linguistically (Rothstein 2010a, Vermonte 2013, Doetjes 2013, Renans 2014), and hence the pattern in (i) is not surprising. We presume that the compatibility of *egy kis* with edible nouns is due to some special characteristics of these nouns, and not to the use of *egy kis* as a determiner.

- (15) *három darab könyv* = (1b)  
 three CL<sub>general</sub> book  
 ‘three books’

This combination of grammatical properties is unexpected in the Chierchian mass/count typology. To account for the optional classifier phenomenon in Hungarian, two approaches have been proposed in the literature: one that is based on the rich classifier phenomena (Csirmaz & Dékány 2014); and another which identifies two possible readings – a mass and a count interpretation – of most nominals in the language (Schvarcz & Rothstein 2017).

According to Csirmaz & Dékány (2014), Hungarian is a language that has classifiers in which all of the nouns are mass, and counting requires either an explicit lexicalized classifier, such as *darab*, or a null sortal classifier (both glossed as CL<sub>general</sub> in (16)). Even though in their paper the term ‘classifier language’ is used in a descriptive sense, i.e. a language that has sortal classifiers, both properties, the existence of classifiers and nouns having a basic mass denotation, make Hungarian a classifier language in the sense of the Chierchian typology.

- (16) a. *három darab könyv*  
 three CL<sub>general</sub> book  
 ‘three books’  
 b. *három  $\emptyset_{CL}$  könyv*  
 three CL<sub>general</sub> book  
 ‘three books’

Even though the Hungarian-as-a-classifier-language approach in Csirmaz & Dékány (2014) explains the usage of sortal classifiers, it does not explain other facts about Hungarian, such as the meaning of the plural marker, or other mass/count-language-like traits. Schvarcz & Rothstein (2017) show that plural marking is sensitive to the mass/count distinction: with notional count nouns, the plural suffix yields a predicate denoting a set of pluralities of individuals, see (17), while with notional mass nouns, pluralization yields an abundance reading (18a), or a plurality of kinds (18b), or a plurality of events reading (18c) (see Corbett 2000, Doron & Müller 2013 and Epstein-Naveh 2015 a.o.).

- (17) *János egész nap fej-ek-et simogat.*  
 János whole day head-PL-ACC stroke.  
 ‘János strokes heads all day.’
- (18) a. *A csatornahálózat nem alkalmas a nagy víz-ek elvezetésé-re.*  
 the sewer.system not suitable the big water-PL diversion-SUBL  
 ‘The sewer system is not suitable for diverting a lot of water.’  
 b. *János összehasonlította a víz-ek ár-á-t.*  
 János compared the water-PL price-POSS-ACC  
 ‘János compared the prices of the different kinds of water.’  
 c. *A jégkorszak végén jöttek a nagy víz-ek.*  
 the ice.age end.of came the big water-PL  
 ‘At the end of the ice age, there were many floods.’

Furthermore, Schvarcz (2014) discusses numerous ways for distinguishing between mass and

count expressions in Hungarian. (For further discussion on these issues, see the Schvarcz (2014) and Schvarcz & Rothstein (2017), and the references therein.) Thus, an alternative solution has been offered by Schvarcz & Rothstein (2017) according to which nouns like *könyv*, ‘book’, are ambiguous between a count and mass interpretation. On the mass reading, the classifier is obligatory (19a), while on the count reading, the classifier is impossible (19b).

- (19) a. *három* \*(*darab*) [*könyv*]<sub>MASS</sub>  
 three CL<sub>general</sub> book  
 ‘three books’
- b. *három* \*(*darab*) [*könyv*]<sub>COUNT</sub>  
 three CL<sub>general</sub> book  
 ‘three books’

Both Csirmaz & Dékány (2014) and Schvarcz & Rothstein (2017) assume that classifiers in Hungarian are required by mass syntax; the role of the classifier is to map mass predicates onto sets of discrete countable units. They both build on the Chierchian framework in which count nouns are not expected to take classifiers: Csirmaz & Dékány (2014) equate Hungarian to a true classifier language; Schvarcz & Rothstein (2017) propose a “mixed system” and two ways of counting: “a single lexical item can be used in counting contexts either as a count noun [and can be directly modified by a numeral] or in a sortal classifier construction” (p. 207). Moreover, both approaches assume that the two possible constructions for expressing the number of some entities – a Num-expression and a CL-expression – are interpreted in the same way.

In this paper we discuss data which suggest a different analysis. We focus on nouns like *könyv* ‘book’, which behave both like mass and count nouns based on their co-occurrence with classifiers. Here we refer to these nouns as *mixed nouns*. We show that mixed nouns can have a different interpretation depending on the kind of expression (Num or CL) they occur in. Based on this data, we claim that the classifier optionality in Hungarian is not sensitive to the grammatical distinction between mass and count nouns. Instead, we argue that classifiers are elements with their own lexical semantic content and their role in Hungarian CL-expressions is not to individuate or divide mass (as mainstream approaches to classifiers suggest), but to determine what type of objects in the denotation of the noun are counted by the numeral.<sup>5</sup>

### 3 The key observation

The main observation of the paper is that in some contexts, the interpretation of Num-expressions and that of CL-expressions may be different despite what the previous approaches suggest. This difference can be observed with mixed nouns that are polysemous having a so-called *physical object* and an *informational object* sense (Pustejovsky 1995, Asher 2011, Gotham 2017, a.o.),

---

<sup>5</sup>The phenomenon of Hungarian CL-expressions alludes to the debate on the semantic content of classifiers. Sortal classifiers have been analyzed as non-lexical semantically redundant elements (Her & Hsieh 2010, Cheng & Sybesma 2012). Other analyses have treated them as determining the denotation of a noun in counting expressions (Bale, Coon & Arcos López 2016, Bale, Coon & Arcos López 2019, and Dékány 2021 for Hungarian). The analysis in this paper aligns with the latter approach. However, we do not wish to provide generalizations about the lexical content of classifiers cross-linguistically, our aim is to compositionally capture the role of the general classifier *darab* in Hungarian.

e.g. *könyv* ‘book’, *festmény* ‘painting’, *magazin* ‘magazine’, etc. If such nouns appear in Num-expressions, the numeral can count the objects of either type in the denotation of the noun – given that the context itself does not specify the interpretation of the noun. However, if the noun occurs in a CL-expressions in the same context, the numeral can only count physical objects. This meaning-shift between Num-expressions and CL-expressions is illustrated in (20).

- (20) a. SCENARIO: Amelia is an adventurous librarian whose main hobby is to look for ancient and medieval books that got lost over the centuries. Last month, she had a particularly successful mission in Alexandria: she found a codex containing three works by Valentina. (Valentina is a medieval author whose work is mostly lost, but well-documented, so it is known that she only wrote books.) There was nothing else in the codex, and Amelia found nothing else last month.
- b. (A *múlt hónap-ban*) Amélia **egy (darab) könyv-et** talált.  
the last month-INE Amelia one (CL<sub>general</sub>) book-ACC found  
‘Last month, Amelia found one book.’
- c. (A *múlt hónap-ban*) Amélia **három (#darab) könyv-et** talált.  
the last month-INE Amelia three (CL<sub>general</sub>) book-ACC found  
‘Last month, Amelia found three books.’

In the scenario described in (20a), there is one book in the *physical object* sense (the codex Amelia found), and there are three books in the *informational object* sense (the works by Valentina in the codex). That is, the noun *könyv*, ‘book’ can be used to refer to books both in the physical and the informational object sense. In this context, both the Num-expression and the CL-expression are felicitous if the numeral in the expression corresponds to the number of physical books (see (20b)), however, only the Num-expression is felicitous if the numeral in the expression corresponds to the number of books in the informational object sense (see (20c)).

For further examples of this meaning-shift see (21) with the noun *festmény* ‘painting’, and (22) with the noun *magazin* ‘magazine’.

- (21) a. SCENARIO: In the art museum, in Room 303 there is only one item on exhibit: a two-sided painting depicting people’s front on the one side, and their backs on the other. The two paintings were painted by two different painters and cannot be separated from each other.
- b. A *303-as terem-ben egy (darab) festmény* van kiállítva.  
the 303 room-INE one (CL<sub>general</sub>) painting is exhibited  
‘In Room 303, there is one painting on exhibit.’
- c. A *303-as terem-ben két (#darab) festmény* van kiállítva.  
the 303 room-INE two (CL<sub>general</sub>) painting is exhibited  
‘In Room 303, there are two paintings on exhibit.’
- (22) a. SCENARIO: Amelia went to the news stand this morning to buy her monthly magazines. She usually buys 3 different ones, but this morning she could only find two of her usual magazines, published as a special joint issue, where one of the magazines is published on one side, and the other magazine on the other side, upside down. Amelia bought nothing else but this special joint issue this morning.
- b. (Ma reggel) Amélia **egy (darab) magazin-t** vett.  
today morning Amelia one (CL<sub>general</sub>) magazine-ACC bought

- ‘This morning, Amelia bought one magazine.’  
 c. (*Ma reggel*) *Amélia két (#darab) magazin-t vett.*  
 today morning Amelia two (CL<sub>general</sub>) magazine-ACC bought  
 ‘This morning, Amelia bought two magazines.’

As the examples in (20)–(22) demonstrate, Num-expressions and CL-constructions do not necessarily have similar interpretations. We summarized the main difference in Table 2.

Expression	What is counted by the Num
Num	physical or informational objects
CL	physical objects

Table 2: The types of individuals counted in different expressions

For this meaning-shift between the two kinds of expressions to occur, the following conditions have to be met: 1. the noun has to be polysemous and have both *physical object* and *informational object* senses, and 2. the context must not restrict the interpretation of the noun to one of its possible senses.<sup>6</sup>

## 4 Problems with the previous approaches

The meaning-shift discussed in the above section poses a problem for previous approaches to mixed nouns and the classifier-phenomenon in Hungarian. As mentioned before, both Csirmaz & Dékány (2014) and Schvarcz & Rothstein (2017) assume that Num-expressions and CL-expressions result in the same interpretation, but they base this assumption on different grounds.

According to Csirmaz & Dékány (2014), the null classifier that is assumed to be present in the structure of Num-expressions, and the overt classifier *darab* (glossed as CL<sub>general</sub> in our examples) are different realizations of the same grammatical element. Furthermore, the same syntactic structure is proposed for Num-expressions and CL-expressions. However, if the two had the same underlying structure, a difference in their interpretation would not be expected. As our data indicate, the two expressions cannot have the same interpretation.

Since the difference in the interpretation of the two expressions correlates with the presence or absence of the classifier, to maintain the assumption about Hungarian being a classifier language, we could slightly modify the approach in Csirmaz & Dékány (2014) and assign the null classifier

---

<sup>6</sup>By *context*, we mean both the linguistic and the extra-linguistic environment of the whole nominal expression. While the extra-linguistic context may specify the interpretation of the noun by providing potential referents for the nominal expression, the linguistic context might involve selectional restrictions imposed by some other expression in the sentence. For example, the verbs *write* or *read* select for informational objects in the denotation of *book* when the noun appears as (part of) the direct object of the verb. Thus, the meaning-shift described above cannot be observed in sentences with these verbs, as shown in (ii).

(ii) *Valentina írt/olvasott három (darab) könyv-et.*  
 Valentia wrote/read three CL<sub>general</sub> book-ACC  
 ‘Valentina wrote/read three books’

and the overt classifier *darab* different interpretations. Dékány (2021) indeed claims so: she observes that the overt classifier is more restrictive regarding the types of available divisions imposed on a mass structure than the null classifier. A clear advantage of such an amendment is that this way we can account for the fact that some nouns in Hungarian cannot occur in Num-expressions, but can co-occur with *darab* (see an example in (23)), which is unexpected if the null classifier and *darab* are merely different realizations of the same grammatical element.

- (23) *három #(darab) márvány*  
 three CL<sub>general</sub> marble  
 ‘three pieces of marble’

However, this modified version of Csirmaz & Dékány (2014) raises a more general question: if Hungarian is a system where classifiers do not have to be realized overtly, why is it that only certain nouns can co-occur with covert classifiers, while others cannot? In other words, given the assumption that classifiers must modify all nouns in order to form grammatically count constructions, there is no obvious explanation why there is a covert classifier available only for some nouns, and there are no other covert classifiers that could potentially modify others.<sup>7</sup> Moreover, even if we assign different semantics for the null classifier and *darab*, by maintaining the assumption that Hungarian is a classifier language, we still cannot account for the facts about Hungarian suggesting that there is a distinction between mass and count nouns, as discussed in Section 2.2.

Schvarcz & Rothstein (2017) presume that in Hungarian the presence or absence of the classifier correlates with the ‘massness’ or ‘countness’ of the noun in the construction. In their analysis, mixed nouns are ambiguous between being mass and count and that is why they can occur both in Num-expressions and CL-expressions. Both expressions are counting expressions, but while in Num-expressions, the countable entities are denoted by the noun itself, in CL-expressions, these countable entities are in the denotation of the noun modified by the classifier.

Even though Schvarcz & Rothstein (2017) could potentially account for differences in the interpretations of the two constructions, the observed pattern we summarized in Table 2 is not a pattern of ambiguity – two mutually exclusive interpretations – but rather one of underspecification, when one interpretation involves the other as an alternative interpretation. Still, one could assume that the difference in the interpretation of Num-expressions and that of CL-expressions is due to the difference in the interpretations of count and the mass versions of the noun. To account for the observed pattern following Schvarcz & Rothstein (2017), we must assume that the difference in the two versions of the noun is as follows: the count version denotes both physical and informational objects, whereas the mass version denotes physical objects only. While this way we could account for the meaning-shift between Num-expressions and CL-expressions, we could not account for the fact that nouns in CL-expressions can be modified by adjectives that apply to the informational objects in the denotation of the noun, as shown in (24). In (24), the adjective *unalmas* ‘boring’ can only apply to informational books, hence the denotation of *könyv* ‘book’ cannot be restricted to physical books. This shows that the meaning-shift between Num-expressions and CL-expressions cannot be reduced to noun ambiguity, thus, the approach in Schvarcz & Rothstein (2017) is not

<sup>7</sup>An additional conceptual argument against amending Csirmaz & Dékány (2014) as described above is that by assuming that the null classifier is different from all overt classifiers semantically, we assign unique semantic content for a grammatical element that by definition cannot be realized phonologically. It is difficult, if not impossible, to test such a theory empirically, as a grammatical structure with the null classifier and an ungrammatical structure without the null classifier are spelled out exactly the same way.

suited to account for the data in (20)–(22).

- (24) *három darab unalmas könyv*  
 three CL<sub>general</sub> boring book  
 ‘three boring books’

We take that the case of mixed nouns and the potential difference in the interpretation of Num-expressions and CL-expressions show that the classifier optionality phenomenon in Hungarian is not sensitive to the grammatical distinction between mass and count nouns. If classifiers in expressions displaying the phenomenon were purely an indicator of some grammatical property of the noun they modify, there should not be any semantic effect associated with their presence or absence in a construction. However, this is not borne out by the data: the interpretation of CL-expressions is more restricted than that of Num-expressions. Therefore, the presence or absence of classifiers in constructions with mixed nouns is affected by semantic factors.

## 5 The analysis of Hungarian classifiers

### 5.1 An MSSC-approach

Before we give our analysis of Hungarian classifiers, we present our assumptions about the semantics of cardinal numerals and counting. Numerals in Hungarian are treated as expressions referring to entities of type  $n$ , i.e. the numerical value expressed by the numeral; see our lexical entry for the numeral *three* in (25). In the prenominal modifier position, numerals are shifted from terms of type  $n$  into predicate modifiers of type  $\langle\langle e, t \rangle, \langle e, t \rangle\rangle$  (see a similar analysis of Hungarian numerals in Schvarcz (under review)).<sup>8</sup> Numerals in prenominal modifier position are devices for the operation of counting: they give the cardinality of an entity in the denotation of the predicate they modify, where the cardinality of an entity is understood as the cardinality of the set containing all the parts of the entity that are the smallest elements in the denotation of the predicate; see (26).<sup>9</sup>

- (25)  $\llbracket \text{three} \rrbracket = 3$

- (26)  $\llbracket \text{three}_{\text{PRENOM}} \rrbracket = \lambda P_{\langle e, t \rangle} . \lambda x_e . P(x) \wedge |x| = 3$ , where  $|x| = 3 \leftrightarrow |\{y : y \leq x \wedge P(y) \wedge \neg \exists z . z < y \wedge P(z)\}| = 3$

Based on our observations summarized in Table 2 above, we propose a purely semantic analysis for Hungarian classifiers according to which the role of classifiers is to specify what kind of entities are counted by the numeral in the construction. In our approach, the counting operation introduced by classifiers selects parts of the entity denoted by the noun that are physically distinct, thereby the complement of a classifier is in fact a count noun, contrary to what has been previously assumed.<sup>10</sup>

<sup>8</sup>We treat numerals in a prenominal modifier position as intersective modifiers, just as in Landman (2003) and Rothstein (2017). However, there these kinds of modifiers are treated as expressions of type  $\langle e, t \rangle$  that combine with other expressions of type  $\langle e, t \rangle$  via a special compositional operation, such as MODIFY in Chung & Ladusaw (2006). Here we assume that intersective modifiers combine with predicates via function application.

<sup>9</sup>There are more fine-grained theories of atomicity and counting (see Rothstein 2010b, Sutton & Filip 2016, a.o.), but the one in (26) is sufficient for the purposes of this paper. Our analysis of Hungarian classifiers should be compatible with these other theories.

Here we undertake the approach that physically distinct entities have the property of being Maximally Strongly Self-Connected (MSSC, Grimm 2012). MSSC-entities are entities that are whole in the topological sense with respect to a property  $P$ , that is, they are the smallest physically bounded entities in the denotation of the predicate. If  $y$  is a Maximally Strongly Self-Connected entity given the predicate  $P$ , the relation MSSC holds between  $y$  and  $P$  (formally represented as  $\text{MSSC}(y, P)$ ). We analyze the general classifier *darab* as incorporating this MSSC-property into the counting operation, see (27).

$$(27) \quad \llbracket \text{darab} \rrbracket = \llbracket \text{CL}_{\text{general}} \rrbracket = \lambda P_{\langle e,t \rangle} . \lambda n_n \lambda x_e . P(x) \wedge |x|_{\text{MSSC}} = n, \text{ where } |x|_{\text{MSSC}} = n \leftrightarrow |\{y : y \leq x \wedge \text{MSSC}(y, P)\}| = n$$

Given our assumptions about the semantics of numerals in (26) and that of the classifier *darab* in (27), we capture the semantic difference between Num-expressions and CL-expressions in Hungarian as follows: Num-expressions are analyzed as predicates that hold of any entity in the denotation of the noun, and this entity has  $n$  number of parts that are atomic with respect to the property expressed by that noun; see (28). We assume that nouns like *könyv* ‘book’, denote both physical and informational objects as atoms (and composite objects made up of physical and informational parts, as in Gotham (2017)), we straightforwardly account for the fact that in Num-expressions entities both in the physical object and the informational object sense can be counted. In contrast, in CL-expressions, the classifier restricts the counting operation to MSSC-entities, thus the numeral provides the number of MSSC-parts of the entity to which the CL-expression applies; see (29).

$$(28) \quad \llbracket (1a) \rrbracket = \llbracket \text{három könyv} \rrbracket = \lambda x . * \text{BOOK}(x) \wedge |x| = 3$$

$$(29) \quad \llbracket (1b) \rrbracket = \llbracket \text{három darab könyv} \rrbracket = \lambda x . * \text{BOOK}(x) \wedge |x|_{\text{MSSC}} = 3$$

As (28) and (29) demonstrate, our analysis of *darab*  $\text{CL}_{\text{general}}$  in (27) can account for the meaning-shift represented in Table 2. Moreover, in order to maintain a unified analysis to all classifiers in Hungarian, the semantics of *darab* can be directly extended to other classifiers like *szem* ‘ $\text{CL}_{\text{small round object}}$ ’ *fej* ‘ $\text{CL}_{\text{big round object}}$ ’ etc, even though, to our knowledge, these classifiers do not combine with nominals that exhibit the informational/physical object polysemy. We assume that these classifiers select nominals based on the size and shape of the entities in their denotation. Thus, they can be analyzed as lexically more specified versions of *darab*  $\text{CL}_{\text{general}}$ , each of them adding some additional properties to be fulfilled by the entities that are counted. As we did with *darab* in (29), we represent these extra properties of these other classifiers as restrictions on the entities selected by the counting operation; see our lexical entry for *szem* in (30).

$$(30) \quad \llbracket \text{szem} \rrbracket = \llbracket \text{CL}_{\text{small round object}} \rrbracket = \\ = \lambda P_{\langle e,t \rangle} . \lambda n_n \lambda x_e . P(x) \wedge |x|_{\text{MSSC-SMALL-ROUND}} = n, \text{ where } \\ |x|_{\text{MSSC-SMALL-ROUND}} = n \leftrightarrow |\{y : y \leq x \wedge \text{MSSC}(y, P) \wedge \text{SMALL}(y) \wedge \text{ROUND}(y)\}| = n$$

---

<sup>10</sup>We acknowledge that this assumption has ramifications for the typological and generative theory of classifiers. In this paper, we do not undertake to explain the theoretical implications of classifiers combining with count nouns, but propose an analysis that treats the Hungarian classifier *darab* as a restrictor on the domain of counting.

## 5.2 Some remarks on mass nouns, abstract nouns and concrete nouns

Before concluding the paper, we would like to highlight some further characteristics of our analysis of Hungarian classifiers. First, since we assume that the presence or absence of the classifier *darab* in Hungarian is not related to the grammatical mass/count distinction, we maintain that the reason why some nouns in Num-expressions are ill-formed is also semantic. That is to say, there are no rules in our grammar that prevent any noun from being directly modified by a numeral. According to our analysis in (26), numerals as prenominal modifiers are sensitive to the non/atomicity of the denotation of the noun they combine with, as they provide the number of parts of an entity that are atomic relative to the property expressed by the noun. Thus, if the noun in a Num-expression is such that its denotation is non-atomic – which is the case in most of the nouns that are grammatically mass – the Num-expression can be assigned no interpretation (putting aside typical packaging contexts, see Jackendoff 1991 and Landman 1991), as shown in (31a). However, if such nouns occur in CL-expressions, the expression is semantically well-formed because the counting operation targets MSSC-parts of the entity in the denotation of the noun, see (32a). In the case of nouns whose denotation is lexically non-atomic, MSSC-entities may be physically distinct pieces of matter (blocks, chunks, nuggets, etc.), and we assume that these entities can only be provided by the context.

- (31) a. #*három agyag*  
           three clay  
       b.  $\llbracket \textit{három agyag} \rrbracket = \lambda x. \text{CLAY}(x) \wedge |x| = 3$
- (32) a. *három darab agyag*  
           three CL<sub>general</sub> clay  
           ‘three pieces of clay’  
       b.  $\llbracket \textit{három darab agyag} \rrbracket = \lambda x. \text{CLAY}(x) \wedge |x|_{\text{MSSC}} = 3$

Second, while our analysis correctly predicts the meaning shift described in Section 3, it does not predict that classifiers can in fact occur with abstract nouns, as shown in (33). Abstract nouns in CL-expressions pose the following problem for our analysis: they are assumed to have abstract objects in their denotation, and since abstract objects do not have an extension in physical space, the MSSC-property imposed by the classifier cannot apply to them. Hence, we predict that classifiers should not appear with abstract nouns, but in fact, they do. This could be because the meaning of the classifier is coerced so that the MSSC-property is not understood along physical dimensions, but along some other relevant dimension(s) specified by the lexical meaning of the noun. In the case of *történet*, ‘story’, in (33), the classifier can restrict the counting operation to some arbitrary structural requirement imposed on complete individual stories, such as having a beginning, discussion and an end. Moreover, both authors share the intuition that the main difference between *három történet* and *három darab történet* in (33) can be captured as follows: while the former can pick out three partially overlapping stories, the latter can only apply to three separate ones. This can be attributed to a strong non-overlap requirement imposed by the classifier (where the overlap relation is understood along the dimension specified by the noun). The analysis for cases as in (33) (and as in (ii) in footnote 6) is subject to future research.

- (33) *három (darab) történet*  
       three CL<sub>general</sub> story

‘three stories’

Third, the assumptions in this paper are based on nouns which are polysemous and have both physical and informational object senses. Our analysis can be extended to concrete nouns – nouns which only denote physical objects. In such cases, we correctly predict that the classifier has no major semantic contribution: even though the Num-expression, *három fotel*, and the CL-expression, *három darab fotel*, in (34) are assigned different interpretations (see (34a) and (34b), respectively), in both expressions, the counting operation can only target entities of the physical object type.<sup>11</sup>

(34) *három (darab) fotel*  
three CL<sub>general</sub> armchair  
‘three armchairs’

- a.  $\llbracket \textit{három fotel} \rrbracket = \lambda x. \text{ARMCHAIR}(x) \wedge |x| = 3$   
b.  $\llbracket \textit{három darab fotel} \rrbracket = \lambda x. \text{ARMCHAIR}(x) \wedge |x|_{\text{MSSC}} = 3$

## 6 Conclusions

In this paper, we presented a formal semantic analysis of Hungarian classifiers, according to which classifiers require the counting operation to select physically distinct entities. We argued for such an analysis based on the case of nouns that display the so-called classifier optionality phenomenon, and that are also polysemous having both a physical object and an informational object sense. We showed that with these nouns, the absence or presence of the classifier correlates with a meaning difference: if such nouns combine with a numeral directly, the numeral can count both the physical objects and the informational objects denoted by the noun. In contrast, if such nouns appear with a classifier, the counting operation is restricted to physical objects. This purely semantic difference between counting expressions with and without a classifiers shows that absence or presence of the classifier in Hungarian is not sensitive to the grammatical mass/count distinction, and so the presence or absence of the classifier does not indicate whether the noun is count or mass. Rather, classifiers are elements with their own lexical content that potentially restricts the type of objects to be counted in the denotation of the noun they modify.

## Acknowledgements

We would like to thank Scott Grimm, Gennaro Chierchia, Yasutada Sudo, Gabi Danon and Louise McNally for their helpful feedback. We also thank Zi Huang for the Chinese data, and two anonymous reviewers for useful comments on earlier drafts of the paper. We are grateful for the audiences of the LingDok23 conference and the NaP2019 Workshop on Number and Plurality for the stimulating discussions. This work was partially supported by the Azrieli Fellowship to Brigitta R. Schvarcz.

---

<sup>11</sup>Here we ignore the possible kind-reading of Num-expressions that disappears in CL-expressions, as discussed in Schvarcz & Nemes (2021). However, our analysis is compatible with the observations therein.

## References

- Asher, Nicholas. 2011. *Lexical Meaning in Context. A Web of Words*. Cambridge: Cambridge University Press.
- Bale, Alan, Jessica Coon & Nicolás Arcos López. 2016. *Counting banana trees in Ch'ol: Exploring the syntax and semantics of sortal classifiers*. Manuscript. <https://ling.auf.net/lingbuzz/003057>.
- Bale, Alan, Jessica Coon & Nicolás Arcos López. 2019. Classifiers, partitions, and measurements: exploring the syntax and semantics of sortal classifiers. *Glossa: a Journal of General Linguistics* 4(1).
- Cheng, Lisa & Rint Sybesma. 2012. Classifiers and DP. *Linguistic Inquiry* 43(4). 634–650.
- Chierchia, Gennaro. 1998. Plurality of mass nouns and the notion of “semantic parameter”. In Susan Rothstein (ed.), *Events and Grammar* (Studies in Linguistics and Philosophy 70), 53–103. Dordrecht: Kluwer.
- Chierchia, Gennaro. 2010. Mass nouns, vagueness and semantic variation. *Synthese* 174(1). 99–149.
- Chung, Sandra & William A. Ladusaw. 2006. Chamorro evidence for compositional asymmetry. *Natural Language Semantics* 14. 325–357.
- Corbett, Grevillea. 2000. *Number*. Cambridge: Cambridge University Press.
- Csirmaz, Anikó & Éva Dékány. 2014. Hungarian is a classifier language. In Raffaele Simone & Francesca Masini (eds.), *Word Classes: Nature, Typology and Representations*, 141–160. Amsterdam: John Benjamins.
- Dékány, Éva. 2021. *The Hungarian Functional Nominal Sequence*. Cham: Springer.
- Doetjes, Jenny. 2013. *Quantity systems and the count/mass distinction*. Paper presented at the Workshop on Countability. 16-17 September 2013, Heinrich-Heine Universität Düsseldorf.
- Doron, Edit & Ana Müller. 2013. The cognitive basis of the mass-count distinction: evidence from bare nouns. In Patricia Cabredo Hofherr & Ann Zribi-Hertz (eds.), *Crosslinguistic studies on noun phrase structure and reference*. Leiden: Brill.
- Epstein-Naveh, Noa. 2015. *Pluralization and Mass Nouns – Can They Go Together in Modern Hebrew?* M.A. Thesis, Bar-Ilan University, Ramat Gan.
- Gil, David. 2013. Numeral classifiers. In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <https://wals.info/chapter/55>.
- Gotham, Matthew. 2017. Composing criteria of individuation in copredication. *Journal of Semantics* 34(2). 333–371.
- Greenberg, Joseph. 1974. *Language Typology: A Historical and Analytic Overview*. Berlin: de Gruyter.
- Grimm, Scott. 2012. Degrees of countability: a mereotopological approach to the mass/count distinction. In Anca Cherches (ed.), *Proceedings of Semantics and Linguistics Theory 22*, 584–603.
- Her, One-Soon & Chen-Tien Hsieh. 2010. On the semantic distinction between classifiers and measure words in Chinese. *Language and Linguistics* 11(3). 527–551.
- Jackendoff, Ray. 1991. Parts and boundaries. *Cognition* 41(1). 9–45.
- Krifka, Manfred. 1989. Nominal reference, temporal constitution and quantification in event semantics. In Renate Bartsch, Johan van Benthem & Peter van Emde Boas (eds.), *Semantics and*

- Contextual Expressions* (Groningen-Amsterdam Studies in Semantics), 75–116. Dordrecht: Foris.
- Krifka, Manfred. 1995. Common nouns: a contrastive analysis of English and Chinese. In Gregory N. Carlson & Francis Jeffry Pelletier (eds.), *The Generic Book*, 398–411. Chicago: Chicago University Press.
- Landman, Fred. 1991. *Structures for Semantics*. Dordrecht: Kluwer.
- Landman, Fred. 2003. Predicate-argument mismatches and the adjectival theory of indefinites. In Martine Coene & Yves D’Hulst (eds.), *From NP to DP, Vol. 1: The Syntax and Semantics of Noun Phrases*, 211–237. Amsterdam: John Benjamins.
- Landman, Fred. 2004. *Indefinites and the Type of Sets*. Oxford: Blackwell.
- Link, Godehard. 1983. The logical analysis of plural and mass terms: a lattice-theoretical approach. In Rainer Bäuerle, Christoph Schwarze & Arnim von Stechow (eds.), *Meaning, Use and Interpretation of Language*, 302–323. Berlin: Walter de Gruyter.
- Pelletier, Francis Jeffry. 1975. Non-singular reference: some preliminaries. *Philosophia* 5(4). 451–465.
- Pustejovsky, James. 1995. *The Generative Lexicon*. Cambridge, MA: MIT Press.
- Renans, Agata. 2014. The semantics of common nouns in Ga (Kwa, Niger-Congo) and their interaction with exclusive particles. In Urtzi Etxeberria, Anamaria Fălăuş, Aritz Irurtzun & Brian Leferman (eds.), *Proceedings of Sinn und Bedeutung 18*, 340–358.
- Rothstein, Susan. 2010a. *Bare nouns semantics, kind interpretations and the Universal Grinder*. Manuscript. Bar-Ilan University, Ramat Gan.
- Rothstein, Susan. 2010b. Counting and the mass count distinction. *Journal of Semantics* 27(3). 343–397.
- Rothstein, Susan. 2017. *Semantics for Counting and Measuring*. Dordrecht: Cambridge University Press.
- Schvarcz, Brigitta R. 2014. *The Hungarians who say -nyi: Issues in Counting and Measuring in Hungarian*. M.A. Thesis, Bar-Ilan University, Ramat Gan.
- Schvarcz, Brigitta R. under review. *Hárman, soka, mindannyian: an event-based analysis to -an/-en marked numerals and quantifiers in Hungarian*. *Journal of Uralic Linguistics*.
- Schvarcz, Brigitta R. & Borbála Nemes. 2021. Classifiers make a difference: kind interpretation and plurality in Hungarian. In Mojmír Dočekal & Marcin Wągiel (eds.), *Formal Approaches to Number in Slavic and Beyond*, 369–396. Berlin: Language Science Press. <https://doi.org/10.5281/zenodo.5082480>.
- Schvarcz, Brigitta R. & Susan Rothstein. 2017. Hungarian classifier constructions, plurality and the mass–count distinction. In Harry van der Hulst & Anikó Lipták (eds.), *Approaches to Hungarian 15. Papers from the 2015 Leiden Conference*, 157–182. Amsterdam: John Benjamins.
- Sutton, Peter R. & Hana Filip. 2016. Counting in context: Count/mass variation and restrictions on coercion in collective artifact nouns. In Mary Moroney, Carol Rose Little, Jacob Collard & Dan Burgdorf (eds.), *Proceedings of Semantics and Linguistics Theory 26*, 350–370.
- Vermonte, Thimotheus. 2013. *Fruits and vegetables in French: A fresh look on systematic mass-count flexibility*. Paper presented at the Workshop on Countability. 16-17 September 2013, Heinrich-Heine Universität Düsseldorf.